Symmetry Breaking in Quantified Boolean Formulae

Gilles Audemard, Said Jabbour, Lahkdar Sais

Symmetry Breaking in Quantified Boolean Formulae

Gilles Audemard* and **Saïd Jabbour** and **Lakhdar Saïs**
CRIL - CNRS, FRE 2499
Université d'Artois - Lens, France
{audemard,jabbour,sais}@cril.univ-artois.fr

Abstract

Many reasoning task and combinatorial problems exhibit symmetries. Exploiting such symmetries has been proved to be very important in reducing search efforts. Breaking symmetries using additional constraints is currently one of the most used approaches. Extending such symmetry breaking techniques to quantified boolean formulae (QBF) is a very challenging task. In this paper, an approach to break symmetries in quantified boolean formulae is proposed. It makes an original use of universally quantified auxiliary variables to generate new symmetry breaking predicates and a new ordering of the QBF prefix is then computed leading to a new equivalent QBF formula with respect to validity. Experimental evaluation of the state-of-the-art QBF solver SEMPROP shows significant improvements (up to several orders of magnitude) on many QBFs instances.

1 Introduction

Solving Quantified Boolean Formulae (QBF) has become an attractive and important research area over the last years. Such increasing interest might be related to different factors, including the fact that many important artificial intelligence (AI) problems (planning, non monotonic reasoning, formal verification, etc.) can be reduced to QBF which is considered as the canonical problem of the PSPACE complexity class. Another important reason comes from the recent impressive progress achieved in the practical resolution of the satisfiability problem. Many solvers for QBFs have been proposed recently (e.g. [Giunchiglia *et al.*, 2001b; Zhang and Malik, 2002; Letz, 2002; Benedetti, 2005]), most of them are obtained by extending satisfiability results. This is not surprising, since QBFs is a natural extension of the satisfiability problem (deciding whether a boolean formula in conjunctive normal form is satisfiable or not), where the variables are universally or existentially quantified.

Some classes of QBFs encoding real-world application and/or AI problems contain many symmetries. Exploiting such structures might lead to dramatically reducing the search space. Symmetries are widely investigated and considered as an important task to deal with the intractability of many combinatorial problems such as constraint satisfaction problems (CSP) and satisfiability of boolean formula (SAT).

A previous work on symmetry breaking predicates for QBF was proposed by [Audemard *et al.*, 2004]. The authors use an hybrid QBF-SAT approach where the set of generated breaking predicates is separated from the QBF formula. Consequently, this approach is solver dependent. Indeed, to solve the hybrid QBF-SAT formula, DPLL-based QBF solvers need to be adapted, whereas, other kind of solvers (e.g. Skizzo) can not be used in this context.

In this paper, we propose a preprocessing approach which is solver independent. Taking as input a QBF with symmetries, we generate a new QBF formula without symmetries equivalent to the original one wrt. validity. To break such symmetries we extend the SAT approach proposed by Crawford [Crawford, 1992; Aloul *et al.*, 2002].

The paper is organized as follows. After some preliminary definitions on quantified boolean formulae, symmetry framework in QBFs is presented. Then a symmetry breaking approach for QBF is described. An experimental validation of our approach is given, showing significant improvements over a wide range of QBF instances. Finally, promising paths for future research are discussed in the conclusion.

2 Technical Background

2.1 Quantified boolean formulae

Let $\mathcal{P}$ be a finite set of propositional variables. Then, $\mathcal{L}_\mathcal{P}$ is the language of quantified Boolean formulae built over $\mathcal{P}$ using ordinary boolean formulae (including propositional constants $\top$ and $\bot$) plus the additional quantification ($\exists$ and $\forall$) over propositional variables.

In this paper, we consider quantified boolean formula Φ in the prenex clausal form $\Phi = Q_k X_k, \dots, Q_1 X_1 \Psi$ (in short $QX\Psi$, QX is called a prefix and Ψ a matrix) where $Q_i \in \{\exists, \forall\}$, $X_k, \dots, X_1$ are disjoint sets of variables and Ψ a boolean formula in conjunctive normal form. Consecutive variables with the same quantifier are grouped. We define $Var(\Phi) = \bigcup_{i \in \{1,\dots,k\}} X_i$ the set of variables of Φ and $VarU(\Phi) = \{x | x \in Var(\Phi),\ x\ is\ universal\}$. A literal is the occurrence of propositional variable in either positive (l) or negative form ($\neg l$). $Lit(\Phi) = \bigcup_{i \in \{1,\dots,k\}} Lit(X_i)$ the set

*supported by ANR "Planevo" project n^{o}JC05_41940.

of complete literals of Φ, where $Lit(X_i) = \{x_i, \neg x_i | x_i \in X_i\}$. The set of literals are encoded using integer numbers i.e. the positive (resp. negative) literal l (resp. $\neg l$) is associated to a positive number α (resp. $-\alpha$). Then, we define $|l|$ as the absolute value of l.

For a given variable $x \in Var(\Phi)$ st. $x \in X_k$, we define $rank(x) = k$. Variables appearing in the same quantifier group are equally ranked.

2.2 Symmetries in Quantified Boolean Formulae

Let $\Phi = Q_1X_1, \dots, Q_mX_m\Psi$ be a QBF and σ a permutation over the literals of Φ i.e. $\sigma : Lit(\Phi) \mapsto Lit(\Phi)$. The permutation σ on Φ is then defined as follows: $\sigma(\Phi) = Q_1\sigma(X_1), \dots, Q_m\sigma(X_m)\sigma(\Psi)$. For example, if Ψ is in clausal form then $\sigma(\Psi) = \{\sigma(c) | c \in \Psi\}$ and $\sigma(c) = \{\sigma(l) | l \in c\}$.

Definition 1 *Let $\Phi = Q_1X_1, \dots, Q_mX_m\Psi$ be a quantified boolean formula and σ a permutation over the literals of Φ. σ is a symmetry of Φ iff*

1. $\forall x \in Lit(\Phi), \sigma(\neg x) = \neg\sigma(x)$
2. $\sigma(\Phi) = \Phi$ *i.e* $\sigma(\Psi) = \Psi$ *and* $\forall i \in \{1, \dots, m\}$ $\sigma(X_i) = X_i$.

Let us note that each symmetry σ of a QBF Φ is also a symmetry of the boolean formula Ψ. The converse is not true. So the set of symmetries of Φ is a subset of the set of symmetries of Ψ.

A symmetry σ can be seen as a list of cycles $(c_1 \dots c_n)$ where each cycle c_i is a list of literals $(l_{i_1} \dots l_{i_{n_i}})$ st. $\forall 1 \leq k < n_i, \sigma_i(l_{i_k}) = l_{i_{k+1}}$ and $\sigma_i(l_{i_{n_i}}) = l_{i_1}$. We define $|\sigma| = \sum_{c_i \in \sigma} |c_i|$ where $|c_i|$ is the number of literals in c_i.

It is well known that breaking all symmetries might lead in the general case to an exponential number of clauses [Crawford *et al.*, 1996]. In this paper, for efficiency and clarity reasons, we only consider symmetries with binary cycles. Our approach can be extended to symmetries with cycles of arbitrary size.

Detecting symmetries of a boolean formula is equivalent to the graph isomorphism problem [Crawford, 1992; Crawford *et al.*, 1996] (i.e. problem of finding a one to one mapping between two graphs G and H). This problem is not yet proved to be NP-Complete, and no polynomial algorithm is known. In our context, we deal with graph automorphism problem (i.e. finding a one to one mapping between G and G) which is a particular case of graph isomorphism. Many programs have been proposed to compute graph automorphism. Let us mention NAUTY [McKay, 1990], one of the most efficient in practice.

Recently, Aloul *et al*. [Aloul *et al.*, 2002] proposed an interesting technique that transforms CNF formula Ψ into a graph G_Ψ where vertices are labeled with colors. Such colored vertices are considered when searching for automorphism on the graph (i.e. vertices with different colors can not be mapped with each others).

In [Audemard *et al.*, 2004], a simple extension to QBFs formulae is given. Such extension is simply obtained by introducing a different color for each set of vertices whose literals belong to the same quantifier group. In this way, literals from different quantifier groups can not be mapped with each others (see the second condition of the definition 1). Then, to detect such symmetries, NAUTY is applied on the graph representation of the QBF.

3 Breaking symmetries in QBFs

Symmetry breaking has been extensively investigated in the context of constraint satisfaction and satisfiability problems. The different approaches proposed to break symmetries can be conveniently classified as dynamic and static schemes. Dynamic breaking generally search and break symmetries using breaking predicates or not [Benhamou and Sais, 1994; Gent and Smith, 2000]. Static breaking schemes refer to techniques that detect and break symmetries in a preprocessing step. For SAT, symmetries are generally broken by generating additional constraints, called symmetry breaking predicates (SBP) [Crawford, 1992; Aloul *et al.*, 2002]. Such SBP eliminates all models from each equivalence class of symmetric models, except one. However, in the general case, the set of symmetry predicates might be of exponential size. In [Aloul *et al.*, 2002], Aloul *et al* extend the approach of Crawford [Crawford, 1992] by using group theory and the concept of non-redundant generators, leading to a considerable reduction in the SBP size.

We briefly recall the symmetry breaking technique introduced by Crawford in [Crawford *et al.*, 1996]. Let Ψ be a CNF formula and $\sigma = \{(x_1, y_1) \dots (x_n, y_n)\}$ a symmetry of Ψ. The SBP associated to σ is defined as follows:

$x_1 \leq y_1$
$(x_1 = y_1) \rightarrow x_2 \leq y_2$
$\dots$
$(x_1 = y_1) \dots (x_{n-1} = y_{n-1}) \rightarrow x_n \leq y_n$

The SBP defined above expresses that, when for all $i \in \{1 \dots k-1\}$ x_i and y_i are equivalent (get the same truth value) and x_k is $true$, then y_k must be assigned to $true$. This reasoning can be extended to QBF provided that the symmetry follows the prefix ordering.

3.1 Motivation

In the following example, we show the main difficulty behind the extension of SAT symmetry breaking predicates (SBP) to QBFs.

Example 1 *Let $\Phi = \forall x_1y_1\exists x_2y_2\ \Psi_1$ be a QBF where $\Psi = (x_1 \vee \neg x_2) \wedge (y_1 \vee \neg y_2) \wedge (\neg x_1 \vee \neg y_1 \vee x_2 \vee y_2)$. The permutation $\sigma_1 = \{(x_1, y_1)(x_2, y_2)\}$ is a symmetry of Φ. Breaking the symmetry σ_1 using the traditional approach, induces the following SBP : $(\neg x_1 \vee y_1) \wedge (\neg x_1 \vee \neg x_2 \vee y_2) \wedge (y_1 \vee \neg x_2 \vee y_2)$. As the clause $(\neg x_1 \vee y_1)$ is universally quantified, the new obtained QBF by adding the SBP to the original QBF leads to an invalid QBF formula.*

To overcome this main drawback, in addition to the classical SBP, new breaking predicates (called $QSBP$) are generated for symmetries containing at least one universal cycle (see definition 3). In such a case, some variables become existentially quantified. These variables will be associated to new additional universally quantified variables. There relationships are expressed in the generated $QSBP$. To safely add such $QSBP$ to the original QBF formula, a new prefix ordering is computed.

After a formal presentation of our approach for a single symmetry, a generalization to arbitrary set of symmetries is then described.

3.2 Breaking a single symmetry

Now, we formally introduce our approach for breaking symmetries in QBF.

Definition 2 *Let $\Phi = Q_1X_1 \dots Q_iX_i \dots Q_mX_m\ \Psi$ be a QBF and σ a symmetry of Φ. We define $\sigma \uparrow X_i$ as the subsequence of the symmetry σ restricted to the cycles involving variables from X_i. Then, the symmetry σ can be rewriten following the prefix ordering as $\{\sigma_1 \dots \sigma_i \dots \sigma_m\}$ such that $\sigma_i = \sigma \uparrow X_i$. When σ respect the prefix ordering, it is called p-ordered.*

In the sequel, symmetries are considered to be p-ordered.

Example 2 *Let $\Phi = \exists x_2y_2\forall x_1y_1\exists x_3y_3\ (\neg x_1 \vee y_1 \vee x_3) \wedge (x_1 \vee \neg x_2 \vee y_3)(x_1 \vee x_2 \vee x_3) \wedge (\neg x_3 \vee \neg y_3) \wedge (x_1 \vee x_2)(x_1 \vee y_2) \wedge (\neg x_1 \vee \neg y_1 \vee \neg x_2 \vee \neg y_2)$ Φ has a symmetry $\sigma = \{(x_1, y_1)(x_2, y_2)(x_3, y_3)\}$. Reordering σ with respect to the prefix leads to $\sigma = \{\sigma_1, \sigma_2, \sigma_3\}$ st. $\sigma_1 = \sigma \uparrow X_1 = (x_2, y_2)$, $\sigma_2 = \sigma \uparrow X_2 = (x_1, y_1)$, $\sigma_3 = \sigma \uparrow X_3 = (x_3, y_3)$.*

Definition 3 *Let Φ be a QBF, σ a symmetry of Φ and $c = (x, y)$ is a cycle of σ. We define x (resp. y) as an in-literal (resp. out-literal). A cycle c is called universal if x and y are universally quantified. A symmetry σ is called universal if it contains at least one universal cycle, otherwise it is called existential.*

For existential symmetries of a QBF, classical SBP [Crawford *et al.*, 1996] can be translated linearly to a CNF formula thanks to new additional variables. The obtained set of clauses can be added to the QBF matrix while preserving its validity. The main problem arises when breaking universal symmetries (see example 1). Indeed, to safely break universal symmetries while keeping the classical SBP, we first reorder the symmetry variables belonging to the same universal quantifier group. This new ordering allows us to determine literals to be likely implied from the SBP. Secondly, as implied universal literals lead to the invalidity of the QBF, an original approach is then proposed to deal with such literals. In the following, the problem behind universal implied literals is illustrated and our approach is then motivated.
Let $\sigma = \{(x_1, y_1), \dots, (x_k, y_k), \dots\}$ be a universal symmetry of a given QBF Φ where (x_k, y_k) is a universal cycle. As mentioned above, σ is ordered according to the prefix of Φ. Suppose x_i and y_i for $1 \leq i \leq k-1$ are assigned the same truth value, if x_k is assigned to $true$ then the universal literal y_k is implied from the SBP. To avoid such a case, the universal quantifier of y_k is substituted with an existential quantifier. However, when x_i and y_i are assigned to different truth values or x_k is assigned to false, y_k must remain universally quantified. To manage these two cases, a new universal variable y'_k is then introduced. This variable plays the same role as y_k in the second case whereas in the first case it becomes useless. The relation between the two variables is expressed using new predicates (called $qsbp(\sigma(y_k))$).

Definition 4 *Let $\sigma = \{c_1 \dots c_k \dots c_n\}$ st. $c_k = (x_k, y_k)$, $1 \leq k \leq n$ be a universal symmetry. We define $QSBP(\sigma) = \cup\{qsbp(\sigma(y_k)), 1 \leq k \leq n\ st.\ y_k\ is\ universal\}$ as the $QSBP$ associated to σ. $QSBP(\sigma)$ is built using the two following steps:*

1. *Adding auxiliary variables : For each universal cycle $c_k = (x_k, y_k) \in \sigma$, we associate a new universal variable y'_k to the out-literal y_k and the universal quantifier of y_k is substituted with an existential quantifier.*
2. *Generating new predicates :*
   - *if (x_1, y_1) is a universal cycle st. $|x_1| \neq |y_1|$ then $qsb(\sigma(y_1)) = \{\neg x_1 \rightarrow (y_1 \leftrightarrow y'_1)\}$*
   - *$\forall k > 1$, if (x_k, y_k) is a universal cycle then $qsbp(\sigma(y_k))$ is made of the following constraints*
     - *$\neg x_k \rightarrow (y_k \leftrightarrow y'_k)$ when $|x_k| \neq |y_k|$*
     - *$((\neg x_j \wedge y_j) \rightarrow (y_k \leftrightarrow y'_k)), \forall j$ st. $1 \leq j < k$*

Example 3 *Let us consider the QBF Φ given in example 1. The symmetry σ of Φ contains one universal cycle (x_1, y_1). Using definition 4, a new variable y'_1 is associated to the variable y_1 which becomes existentially quantified. Then QSBP(σ)= qsbp($\sigma(y_1)$)= $(\neg x_1 \rightarrow (y_1 \leftrightarrow y'_1))$*

As described above, to generate the QSBP new variables are introduced. In the sequel, we describe how such variables are integrated in the QBF prefix.

Definition 5 *Let $\Phi = Q_1X_1 \dots Q_mX_m\ \Psi$ be a QBF, and $\sigma = \{\sigma_1 \dots \sigma_j \dots \sigma_m\}$ be a universal symmetry. Let j st. $Q_j = \forall$ and $\sigma_j = \sigma \uparrow X_j = \{(x_1, y_1) \dots (x_n, y_n)\}$ and $Y' = \{y'_1 \dots y'_n\}$ the set of new variables associated to $\{y_1 \dots y_n\}$ respectively. We define the new ordering of $Var(\sigma_j) \cup Y'$ as follows : $rank(x_k) < rank(y'_k) < rank(y_k)$ st. $1 \leq k \leq n$.*

We define $\mathcal{G}_{\sigma_j}(\mathcal{V}, \mathcal{A})$ as the precedence graph associated to σ_j with $\mathcal{V} = \bigcup_{1\leq k\leq n}\{x_k, y_k, y'_k\}$ and $\mathcal{A} = \{\bigcup_{1\leq k\leq n}\{(x_k, y'_k), (y'_k, y_k)\}\}$.

To rewrite the quantifier group Q_jX_j of the QBF formula (see definition 6), a new ordering is derived by applying topological sort algorithm on the precedence graph. Let us note that such a graph is acyclic. In figure 1, the graph representation of $\sigma_j = \sigma \uparrow X_j$ is illustrated. The ordering $x_1 \dots x_n$ $y'_1 \dots y'_n y_1 \dots y_n$ is then considered.

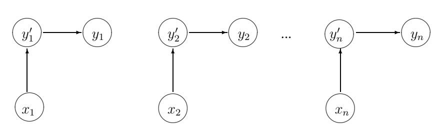

Figure 1: Graph representation of σ_j

Definition 6 *Let $\Phi = Q_1X_1 \dots Q_mX_m\ \Psi$ be a QBF, $\sigma = \{\sigma_1 \dots \sigma_m\}$ a universal symmetry and for $j \in \{1 \dots m\}$ $\sigma_j = \sigma \uparrow X_j = \{(x_1, y_1), \dots (x_n, y_n)\}$. Every quantifier group Q_jX_j of Φ is rewritten :*

- *if $Q_j = \forall$ and $\sigma \uparrow X_j \neq \emptyset$ then*
  $Q^p_jX^p_j = \forall(X_j \backslash Var(\sigma_j))x_1 \dots x_n \exists\alpha\forall y'_1 \dots y'_n \exists y_1 \dots y_n$

(see figure 1) st. $X_j^p = X_j \cup \{y'_1 \dots y'_n\}$ *and* α *is a new variable.*

- *otherwise* $Q_j^p X_j^p = Q_j X_j$

Φ *is then rewritten as*
$\Phi^S = Q_1^p X_1^p \dots Q_i^p X_i^p \dots Q_m^p X_m^p \ \Psi \wedge SBP \wedge QSBP$

Remark 1 *In definition 6, a new variable* α *is introduced to constrain the set of variables* $\{x_1, \dots x_n\}$ *to be assigned before* $\{y'_1, \dots, y'_n\}$.

Property 1 *Let* Φ *be a QBF and* σ *a symmetry of* Φ*, then* Φ *is valid iff* Φ^S *is valid.*

To sketch the proof of property 1, we consider the QBF given in the example 1. Applying our approach on Φ, we obtain the QBF :

$\Phi^s = \forall x_1 \exists \alpha \forall y'_1 \exists y_1 x_2 y_2 \ \Psi_1 \wedge (\neg x_1 \vee y_1) \wedge (\neg x_1 \vee \neg x_2 \vee y_2) \wedge (y_1 \vee \neg x_2 \neg y_2) \wedge (x_1 \vee y_1 \vee \neg y'_1) \wedge (x_1 \vee \neg y_1 \vee y'_1)$
When x_1 is assigned to the value $true$, from the clause $(\neg x_1 \vee y_1)$ we deduce that y_1 is $true$. If x_1 is assigned the value $false$, the original universal variable y_1 is deduced by substitution (y'_1 and y_1 are equivalent) thanks to the added constraint $\neg x_1 \rightarrow (y_1 \leftrightarrow y'_1)$. As $rank(y'_1) < rank(y_1)$, we only need to substitute all occurrences of y_1 by y'_1.

3.3 Breaking all symmetries

Generating QSBP

When considering several symmetries, we can not eliminate them independently by processing each single symmetry using the single symmetry breaking approach described in section 3.2. One needs to consider the interactions between the different symmetries. Indeed, considering an universal out-literal y_k and its associated new variable y'_k, the $qsbp(\sigma(y_k))$ express the conditions under which such literals y_k and y'_k are equivalent. As only one variable y'_k is introduced for each universal out-literal y_k, when y_k appears in several symmetries, the different conditions leading to such equivalence need to be combined.

Definition 7 *Let* Φ *be a QBF, and* $\sigma = \{(x_1, y_1) \dots (x_i, y) \dots\}$. $\sigma' = \{(z_1, w_1) \dots (z_j, y) \dots\}$ *two symmetries of* Φ *with* y *an universal out-literal. The qsbp's associated to* y *with respect to* σ *and* σ' *can be written in the following form (see definition 4) :*

- $qsbp(\sigma(y)) = \{\alpha_1 \rightarrow (y \leftrightarrow y') \dots \alpha_N \rightarrow (y \leftrightarrow y')\}$
- $qsbp(\sigma'(y)) = \{\beta_1 \rightarrow (y \leftrightarrow y') \dots \beta_M \rightarrow (y \leftrightarrow y')\}$.

We define a binary correlation operator η *between* σ *and* σ' *as follows :*
$\eta(\sigma(y), \sigma'(y)) = \{(\alpha_1 \wedge \beta_1) \rightarrow (y \leftrightarrow y') \dots (\alpha_1 \wedge \beta_M) \rightarrow (y \leftrightarrow y') \dots (\alpha_N \wedge \beta_1) \rightarrow (y \leftrightarrow y') \dots (\alpha_N \wedge \beta_M) \rightarrow (y \leftrightarrow y')\}$.

Definition 8 *Let* $\mathcal{S} = \{\sigma^1 \dots \sigma^n\}$ *be a set of symmetries of a given QBF and* y *is universal literal. We define,*

- $\mathcal{S}[y] = \{\sigma^j | \exists (x, y) \in \sigma^j\} = \{\sigma'^1 \dots \sigma'^{|\mathcal{S}[y]|}\}$.
- $qsbp(\mathcal{S}[y]) = \eta(\eta \dots \eta(\sigma'^1, \sigma'^2) \dots, \sigma'^{|\mathcal{S}[y]|}) \dots)$.

Definition 9 *Let* $\mathcal{S} = \{\sigma^1 \dots \sigma^n\}$ *be the set of symmetries of a given QBF* $\Phi = QX \ \Psi$*. The new QBF matrix* Ψ^S *is defined as follows:*

$$\Psi \wedge \left(\bigwedge_{1 \leq i \leq n} SBP(\sigma^i) \right) \wedge \left(\bigwedge_{x \in VarU(\Phi)} qsbp(\mathcal{S}[x]) \right)$$

Prefix ordering

Let us now show how a new QBF prefix is built when considering a set of symmetries (for a single symmetry see definition 6). For a set of symmetries $\{\sigma^1, \dots, \sigma^n\}$, we consider for each universal quantifier group $Q_k X_k$, all the projections $\sigma^i \uparrow X_k$ for each symmetry σ^i. The new quantifier group $Q_k^p X_k^p$ is obtained from the precedence graph representation of all these projections. Let us recall that each symmetry σ^i is considered p-ordered. Additionally, to avoid cycles from the graph representation of $\{\sigma^i \uparrow X_k | 1 \leq i \leq n\}$, each projection $\sigma^i \uparrow X_k$ is considered lexicographically ordered.

Definition 10 *Let* $\sigma = \{(x_1, y_1), \dots, (x_n, y_n)\}$ *be a set of cycles.* σ *is called lexicographically ordered (lex-ordered in short), iff* $x_i > 0$, $x_i \leq |y_i|$ $1 \leq i \leq n$ *and* $x_i < x_{i+1}$ $1 \leq i < n$

Example 4 *Let* $\Phi = Q_1 X_1 \dots Q_m X_m \ \Psi$ *be a QBF,* σ *and* σ' *two symmetries of* Φ *st.* $\sigma \uparrow X_1 = \{(x_2, x_1)(\neg x_3, \neg x_4)(x_5, x_6)\}$, $\sigma' \uparrow X_1 = \{(x_3, x_1)(x_2, \neg x_6)\}$ *and* $Q_1 = \forall$. *With respect to lexicographical order,* $\sigma \uparrow X_1 = \{(x_1, x_2)(x_3, x_4)(x_5, x_6)\}$ *and* $\sigma' \uparrow X_1 = \{(x_1, x_3)(x_2, \neg x_6)\}$ *Figure 2 shows the precedence graph representation of both* $\sigma \uparrow X_1$ *and* $\sigma' \uparrow X_1$. *Applying topological sort algorithm, the quantifier* $Q_1^P X_1^P$ *is then rewritten as* $Q_1^P X_1^P = \forall x_1 x_5 \exists \alpha \forall x'_2 x'_3 \exists x_2 x_3 \forall x'_4 x'_6 \exists x_4 x_6$

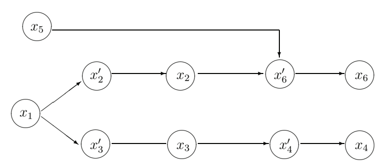

Figure 2: Precedence graph of $\sigma \uparrow X_1$ and $\sigma' \uparrow X_1$

Dealing with universal cycles of the form (y, x) and $(z, \neg x)$

To preserve the equivalence (wrt. validity) between the original QBF and the new generated one, let us now address the last problem arising from the interactions between the different symmetries. As illustrated in the following example, the problem arises for symmetries where a universally quantified out-literal x appears both positively (x) and negatively ($\neg x$).

Example 5 *Let* $\Phi = Q_1 X_1 \dots Q_i X_i \dots Q_m X_m \Psi$ *be a QBF,* σ *and* σ' *two symmetries of* Φ *st.* $\sigma \uparrow X_i = \{(x_1, y_2)\}$ *and* $\sigma' \uparrow X_i = \{(x_2, \neg y_2)\}, Q_i = \forall$.

Using our approach the quantifier group $Q_i X_i$ is rewritten as $\forall (X_i \backslash Var(\sigma, \sigma')) \forall x_1 x_2 \exists \alpha \forall y'_2 \exists y_2$. The generated SBP for σ contains the clause $c = (\neg x_1 \vee x_4)$. For σ' its corresponding SBP contains the clause $c' = (\neg x_2 \vee \neg y_2)$. As the universal variable x_4 is substituted with an existential one, applying the Q-$resolution$ rule [Kleine-Büning *et al.*, 1995] on c and c' leads to a universally quantified resolvent $r = (\neg x_1 \vee \neg x_2)$. Consequently, the new obtained QBF is invalid. Finding a new ordering which avoid this problem is a very challenging task. In example 5, such a problem can be avoided by considering the new ordering $x_1 < y_2 < x_2$ instead of the used lex-ordering. Another possible solution, actually under investigation is to apply composition between symmetries. More precisely, if we consider $\sigma \circ \sigma' \circ \sigma$ we obtain a new symmetry $\sigma'' = (x_1, \neg x_2)$. If in addition to σ and σ', we also consider σ'', then the previous resolvent generated using $Q\text{-}resolution$ is now not universally quantified. Indeed, as x_2 is an out-literal, its universal quantifier is substituted with an existential one. Then the resolvent r contains a literal $\neg x_2$ whose associated variable is existentially quantified. Finally, the quantifier group $Q_i X_i$ can be rewritten as $Q_i^p X_i^p = \forall x_1 \exists \alpha \forall x_2' \exists x_2 \forall y_2' \exists y_2$.

The above discussion gives us an idea on how to solve in the general case, the problem arising from symmetries with universal cycles of the form (y, x) and $(z, \neg x)$. In this paper, such a problem is simply avoided using the following restriction :

Definition 11 *Let $\Phi = Q_1 X_1 \ldots Q_m X_m \Psi$ be a QBF, $y \in VarU(\Phi)$ and $\mathcal{S}$ the set of symmetries of Φ. We define $\mathcal{S}[y] \downarrow y = \{\sigma \downarrow y | \sigma \in \mathcal{S}[y]\}$. For $\sigma = \{(x_1, y_1), \ldots, (x_k, y_k) \ldots, (x_n, y_n)\}$, we define $\sigma\downarrow y_k = \{(x_1, y_1), \ldots, (x_{k-1}, y_{k-1})\}$. The restriction of $\mathcal{S}$ wrt. y is defined as $rt(S, y) = \{\sigma | \sigma \in S, \sigma[y] = \emptyset, \sigma[\neg y] = \emptyset\} \cup \mathcal{S}[y] \cup \mathcal{S}[\neg y]\downarrow\{\neg y\}$. For the set of variables $VarU(\Phi) = \{v_1, \ldots, v_{|VarU(\Phi)|}\}$, we define $rt(S, VarU(\Phi)) = rt(\ldots rt(S, v_1), v_2) \ldots v_{|VarU(\Phi)|}) \ldots)$*

Note that if $S[y] = \emptyset$ or $S[\neg y] = \emptyset$, then $rt(S, y) = S$. Naturally, for a given set of symmetries S, the new QBF formula is generated using $rt(S, VarU(\Phi))$. In this way the obtained formula is equivalent wrt. validity to the original one.

Complexity

Let σ be a symmetry of a QBF and $CNF(QSBP(\sigma))$ the CNF representation of $QSBP(\sigma)$. The worst case spacial complexity of $CNF(QSBP(\sigma))$ is in $O(|\sigma|^2)$. Considering $\sigma = \{\sigma_1, \ldots, \sigma_n\}$ with $\sigma_i = (x_i, y_i)$, $1 \leq i \leq n$, the worst case is reached when all cycles of σ are universal. In this case, $QSBP(\sigma) = \bigcup_{1 \leq i \leq n} qsbp(\sigma_i(y_i))$. $|QSBP(\sigma)| = \sum_{1 \leq i \leq n} |qsbp(\sigma(y_i))|$. $|qsbp(\sigma(y_i))|$ is equal to $2(i-1) + 2 = 2i$ (see definition 4). Then, $|QSBP(\sigma)| = \sum_{1 \leq i \leq n} 2i$ which is equal to $n(n+1)$. More interestingly, using the same new variables introduced for $CNF(SBP)$, the size of $QSBP(\sigma)$ becomes $linear$. Unfortunately, because of the correlations between different symmetries, the $QSBP$ associated to a set of symmetries is exponential in the worst case. In practice, on all the considered QBF instances, the number of applied correlations (definition 7) does not exceed 3. Consequently, the QSBP is most often of reasonable size.

4 Experiments

The experimental results reported in this section are obtained on a Xeon 3.2 GHz (2 GB RAM) and performed on a large panel of symmetric instances (619) available from [Giunchiglia *et al.*, 2001a]. This set of QBF instances contains different families like `toilet`, `k_*`, `FPGA`, `qshifter`. As a comparison, we run the state-of-the-art DPLL-like solver SEMPROP [Letz, 2002] on QBF instances with and without breaking symmetries. The time limit is fixed to 900 seconds. Results are reported in seconds. The symmetry computation time (including detection and QSBP generation) is not reported (less than one second in most cases).

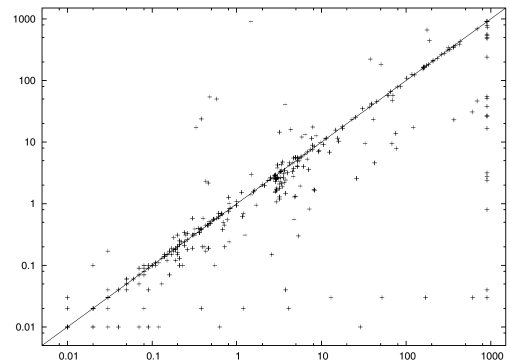

Figure 3: SEMPROP comparison

The scatter plot (in log scale) given in figure 3 illustrate the comprative results of SEMPROP [Letz, 2002] on each QBF instance Φ and Φ^S respectively. The x-axis (resp. y-axis) corresponds to the cpu time tx (resp. ty) obtained by SEMPROP on Φ (resp. Φ^S). Each dot with (tx, ty) coordinates, corresponds to a QBF instance. Dots above (resp. below) the diagonal indicate instances where the original formula Φ is solved faster i.e. $tx < ty$ (resp. slower i.e. $tx > ty$) than the QBF formula Φ^S.

Figure 3 clearly shows the computational gain obtained using symmetry breaking predicates (about 167 instances are solved more efficiently). In some cases the gain is up to 2 orders of magnitude. Of course, there exists some instances where breaking symmetries decreases the performances of SEMPROP(about 50 instances). On the remaining instances, the performance of the solver remains the same with or without breaking symmetries.

Table 1 provides more detailed results on the different QBF families. The second column (NB) represents the number of instances in each family. The third column (U) indicates if the instances contain universal symmetries (Y) or not (N). For each family, S and TT represents the total number of solved instances and the total run-time needed for solving all the instances (900 seconds are added for each unsolved one) respectively.

As we can see, table 1 gives us more comprehensive results with respect to each family. First, breaking symmetries significantly improves SEMPROP performances on many QBF families leading to more solved instances (18 instances). Secondly, the existence of universal symmetries seems to be an important factor for reducing the search time. Not surprisingly, we have also noticed that symmetries between literals occurring in the innermost quantifier group are useless. Indeed, such symmetries does not lead to a great reduction in

			Φ		Φ^S	
family	NB	U	S	TT	S	TT
fpga	8	Y	6	1834	**7**	**921**
blackbox	23	Y	1	19801	**8**	**13668**
scholl	32	Y	17	13687	**18**	**13595**
toilet_c	53	Y	51	2656	**53**	**49**
k_path	40	Y	28	12915	**34**	**7999**
qshifter	6	Y	6	67	6	**61**
tipdiam	76	Y	30	41455	30	41459
asp	104	Y	104	**789**	104	1909
TOILET	7	Y	6	988	6	**926**
term1	6	Y	6	**174**	6	177
strategic	100	N	86	13482	86	**13477**
k_branch	42	N	21	20096	21	**20069**
k_lin	21	N	5	14553	5	14551
k_grz	37	N	23	15242	**24**	**14997**
k_poly	42	N	42	**2031**	42	2068
toilet_a	22	N	22	**12**	22	20
TOTAL	619		454	159789	**472**	**145940**

Table 1: Results on different QBF families

the search tree, since their corresponding variables are assigned last i.e. the formula is considerably reduced by the previous assignments. Finally, for QBF families containing only existential symmetries, breaking them do not improves the search time. On k_branch, k_lin, $toilet\_a$ families containing only existential symmetries, no improvement is observed. Most of these instances correspond to the dots near the diagonal (see figure 3).

5 Conclusion

In this paper, a new approach to break symmetries in QBF formulae is proposed. Using universally quantified auxiliary variables, new symmetry breaking predicates are generated and safely added to the QBF formula. Experimental results show that breaking symmetries leads to significant improvements of the state-of-the-art QBF solver SEMPROP on many classes of QBF instances. These experimental results suggest that for QBF instances containing universal symmetries, significant improvements are obtained. As future works, we plan to investigate the problem arising from the presence of both positive and negative out-literal in the set of symmetries.

References

[Aloul *et al.*, 2002] F. Aloul, A. Ramani, I. Markov, and K. Sakallah. Solving difficult instances of boolean satisfiability in the presence of symmetry. Technical report, University of Michigan, 2002.

[Audemard *et al.*, 2004] G. Audemard, B. Mazure, and L. Sais. Dealing with symmetries in quantified boolean formulas. In *proceedings of SAT*, 2004.

[Benedetti, 2005] M. Benedetti. sKizzo: a Suite to Evaluate and Certify QBFs. In *Proc. of CADE*, 2005.

[Benhamou and Sais, 1994] B. Benhamou and L. Sais. Tractability through symmetries in propositional calculus. *Journal of Automated Reasoning*, 12(1):89–102, 1994.

[Crawford *et al.*, 1996] J. Crawford, M. Ginsberg, E. Luck, and A. Roy. Symmetry-breaking predicates for search problems. In *proceedings of KR*, pages 148–159. 1996.

[Crawford, 1992] J. Crawford. A theorical analysis of reasoning by symmetry in first order logic. In *Proceedings of Workshop on Tractable Reasonning, AAAI*, pages 17–22, 1992.

[Gent and Smith, 2000] I. Gent and B. Smith. Symmetry breaking in constraint programming. In *Proceedings of ECAI*, pages 599–603, 2000.

[Giunchiglia *et al.*, 2001a] E. Giunchiglia, M. Narizzano, and A. Tacchella. Quantified Boolean Formulas satisfiability library (QBFLIB), 2001.

[Giunchiglia *et al.*, 2001b] E. Giunchiglia, M. Narizzano, and A. Tacchella. QuBE : A system for deciding Quantified Boolean Formulas Satisfiability. In *Proceedings of IJCAR*, 2001.

[Kleine-Büning *et al.*, 1995] H. Kleine-Büning, M. Karpinski, and A. Flögel. Resolution for quantified boolean formulas. *Information and computation*, 117(1):12–18, 1995.

[Letz, 2002] R. Letz. Lemma and model caching in decision procedures for quantified boolean formulas. In *Proceedings of Tableaux*, pages 160–175, 2002.

[McKay, 1990] B. McKay. nauty user's guide (version 1.5). Technical report, 1990.

[Zhang and Malik, 2002] L. Zhang and S. Malik. Towards a symmetric treatment of satisfaction and conflicts in quantified boolean formula evaluation. In *Proceedings of the CP*, pages 200–215, 2002.